

What Is Claimed Is:

1. A method for the assignment of phonemes to a lexicon of words, comprising:
using a dynamic time warping algorithm to phonetically transcribe the words by
assigning phoneme sequences to grapheme sequences of the words.

2. A method for the assignment of phonemes to the graphemes producing them
in a lexicon having words (grapheme sequences) and their associated phonetic transcription
(phoneme sequences), comprising:

a) in a first step, determining relative frequency with which phonemes and
graphemes are assigned to one another for each assignment of phonemes and graphemes,

b) in a second step, creating for each word of the lexicon a two-dimensional
matrix (incidence matrix), one index of which is given by the grapheme of the word, and the
second index of which is given by the phoneme of the word,

c) selecting the relative frequencies belonging to the respective
phoneme-grapheme pair and determined in the first step as entries of the matrix,

d) in a third step, logically combining each matrix entry with the aid of a
mathematical operation with the extreme value of the following three preceding matrix
entries:

the entry for the same phoneme and the preceding grapheme in the
word,

the entry for the preceding phoneme and the same grapheme in the
word, and

the entry for the preceding phoneme and the preceding grapheme in
the word,

e) using the first grapheme and the first phoneme of the word as the starting
point in the mathematical operation, and using the modified entries of the matrix in
determining the extreme values, the modified entries being respectively yielded from the
mathematical operation,

f) determining which of the three preceding matrix entries was extreme to
thereby determine a step direction for this matrix entry,

g) in a fourth step, defining the step direction determined for the matrix entry,
starting from the matrix entry for the last phoneme and the last grapheme, and proceeding
along a path through the matrix up to the matrix entry for the first phoneme and the first
grapheme, and

09943091 083401
T07E90 T6D466

h) using the matrix elements along the path to define the assignment of graphemes to phonemes of the word.

3. The method as claimed in claim 2, wherein the relative frequencies in the first step are determined by selecting words from the lexicon in the case of which the number of the graphemes and the number of the phonemes coincide, for the selected words, the graphemes and phonemes are assigned to one another in the sequence of the specification of their graphemes and phonemes in the lexicon.

4. The method as claimed in claim 1, wherein after execution of the assignment of graphemes to phonemes for each word of the lexicon, these assignments are used to determine the position-dependent relative frequency with which at least one of the following combination occur:

- a phoneme produced by two or more graphemes,
- two or more phonemes produced by a grapheme,
- two or more graphemes assigned to a phoneme, and
- a grapheme assigned to two or more phonemes.

5. The method as claimed in claim 4, wherein the assignment of graphemes to phonemes within a word is corrected with the aid of the position-dependent relative frequencies.

6. The method as claimed in claim 3, wherein after execution of the assignment of graphemes to phonemes for each word of the lexicon, these assignments are used to determine the position-dependent relative frequency with which at least one of the following combinations occur:

- a phoneme produced by two or more graphemes,
- two or more phonemes produced by a grapheme,
- two or more graphemes assigned to a phoneme, and
- a grapheme assigned to two or more phonemes.

7. The method as claimed in claim 6, wherein the assignment of graphemes to phonemes within a word is corrected with the aid of the position-dependent relative frequencies.

8. The method as claimed in claim 7, wherein
after assigning graphemes to phonemes for selected words in the sequence of the specification, for each word of the lexicon, the corrected assignments are used to recalculate the position-dependent relative frequency with which a phoneme is produced by two or more graphemes, or two or more phonemes are produced by a grapheme; and
the recalculated position dependent relative frequencies are used to again assign graphemes to phonemes for selected words in the sequence of the specification.

9. The method as claimed in claim 8, wherein in order to determine the relative frequencies, only those assignments are taken into account in which the matrix entry for the last phoneme and the last grapheme exceeds a prescribed threshold value after execution of the multiplications.

10. The method as claimed in claim 2, wherein
the matrix entry for the first phoneme and the first grapheme of each word is set to 1;
the matrix entry for the last phoneme and the last grapheme of each word is set to 1;
the matrix entry for the first phoneme and the last grapheme of each word is set to 0;
and
the matrix entry of the last phoneme and the first grapheme of each word is set to 0.

11. The method as claimed in claim 2, wherein if in the determination of the maximum value of the three preceding matrix entries the matrix entry for the preceding phoneme and the preceding grapheme in the word and one of the other two entries are of equal magnitude, the matrix entry for the preceding phoneme and the preceding grapheme in the word is regarded as a maximum.

12. A computer system for the assignment of phonemes to a lexicon of words, comprising:
a storage device for storing a computer program on a storage medium; and
a processing unit for loading the computer program from the storage device and for executing the computer program so as to use a dynamic time warping algorithm to phonetically transcribe the words by assigning phoneme sequences to grapheme sequences of the words.

13. A computer readable medium storing a program for controlling a computer to perform a method for the assignment of phonemes to the graphemes producing them in a lexicon having words (grapheme sequences) and their associated phonetic transcription (phoneme sequences), comprising:

- a) in a first step, a determining relative frequency with which phonemes and graphemes are assigned to one another for each assignment of phonemes and graphemes,
- b) in a second step, creating for each word of the lexicon a two-dimensional matrix (incidence matrix), one index of which is given by the grapheme of the word, and the second index of which is given by the phoneme of the word,
- c) selecting the relative frequencies belonging to the respective phoneme-grapheme pair and determined in the first step as entries of the matrix,
- d) in a third step, logically combining each matrix entry with the aid of a mathematical operation with the extreme value of the following three preceding matrix entries:
 - the entry for the same phoneme and the preceding grapheme in the word,
 - the entry for the preceding phoneme and the same grapheme in the word, and
 - the entry for the preceding phoneme and the preceding grapheme in the word,
- e) using the first grapheme and the first phoneme of the word as the starting point in the mathematical operation, and using the modified entries of the matrix in determining the extreme values, the modified entries being respectively yielded from the mathematical operation,
- f) determining which of the three preceding matrix entries was extreme to thereby determine a step direction for this matrix entry,
- g) in a fourth step, defining the step direction determined for the matrix entry, starting from the matrix entry for the last phoneme and the last grapheme, and proceeding along a path through the matrix up to the matrix entry for the first phoneme and the first grapheme, and
- h) using the matrix elements along the path to define the assignment of graphemes to phonemes of the word.